



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|--|-------------|----------------------|---------------------|------------------|
| 09/369,031 | 08/04/1999 | LEX S. OLORENSHAW | 3188.01/1144 | 6238 |
| 24272 | 7590 | 12/17/2003 | EXAMINER | |
| Gregory J. Koerner Simon & Koerner LLP 10052 Pasadena Avenue, Suite B Cupertino, CA 95014 | | | KNEPPER, DAVID D | |
| | | | ART UNIT | PAPER NUMBER |
| | | | 2654 | |
| DATE MAILED: 12/17/2003 | | | | |

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) |
|------------------------------|------------------------|---------------------|
| | 09/369,031 | OLORENSHAW ET AL. |
| | Examiner | Art Unit |
| | David D. Knepper | 2654 |

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 11 August 2003.

2a) This action is **FINAL**. 2b) This action is non-final.

3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-5,7-29 and 31-60 is/are pending in the application.
4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) Claim(s) _____ is/are allowed.

6) Claim(s) 1-5,7,9-19,21-29,31,33-43,45-51,53,55 and 56 is/are rejected.

7) Claim(s) 8,20,32,44,52,54 and 57-60 is/are objected to.

8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.

10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.

 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) All b) Some * c) None of:
1. Certified copies of the priority documents have been received.
2. Certified copies of the priority documents have been received in Application No. _____.
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
* See the attached detailed Office action for a list of the certified copies not received.

13) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.
a) The translation of the foreign language provisional application has been received.

14) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) Notice of References Cited (PTO-892) 4) Interview Summary (PTO-413) Paper No(s). ____ .
2) Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) Notice of Informal Patent Application (PTO-152)
3) Information Disclosure Statement(s) (PTO-1449) Paper No(s) 6) Other: _____

1. Applicants' amendment (paper #11) received 11 August 2003 has been considered. Claims 1-5, 7-29, and 31-60 are pending. Claims 6 and 30 have been canceled.

Specification

2. The specification is objected to because of improper incorporations by reference to 3 non-patent documents on pages 8, 10 and 13. The relevant information from these documents must be inserted into the specification. They are over 10 years old (dated 1989 and 1982), and, as such, are admitted prior art. The applicant is referred to section 609 of the MPEP, which discusses treatment of materials incorporated by reference.

*NOTE => The applicant's last Amendment was non-responsive to the above objection informing the USPTO that the applicant wished to defer responding until claims are allowed. The applicant is required to show the legal reference (Rule or Law) that requires such deference by the USPTO. Failure to provide such legal reference will result in treatment of this argument to the objection as a deliberate failure to respond and could result in abandonment of the application.

Claims

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary

skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. Claims 1-5, 7, 9-19, 21-29, 31, 33-43, 45-51, 53, 55, 56 are rejected under 35 U.S.C. § 103 as being unpatentable over Kao (6,316,712) in view of Gillick (5,715,367).

As per claim 1, “Speech processing” is taught by Kao’s phonetic modeling using acoustic decision tree, title:

“speech data generated from one or more speech sources” (English has about 40 to 50 phones . . . phonetic models can be clustered to not just reduce the number of models but also increase the training robustness, co. 1, lines 8-32);

“an enhanced phone set that includes acoustic-phonetic symbols and connectors for extending said enhanced phone set” (his few examples in col. 2, lines 27-40, symbols and connectors are shown in columns 4-5 and columns 8-9 which show example symbols marking phone models and transitions between them -- some specific phones and interconnecting relationships are shown in Gillicks’ figure 9);

“transcription generated by a transcription process that selects appropriate phones from said enhanced phone set to represent said speech data” (suggested by his use of symbols in col. 2, lines 30-40 – see unique transcription symbols used by Gillick in column 28-29).

Kao uses the term “transcription” in column 8, lines 44-46 indicating that the sequence of phones [are supervised] according to the transcription provided with the corpus. This includes inter-word context. Thus, he teaches that his models are used to recognize speech and this requires using the models to compare input speech and transcribe or otherwise phone symbols

element by element to form words. Gillick shows some specific transcription symbols in columns 28-29. The process of “transcription” is considered equivalent to Gillick’s use of phonetic spelling, column 2, lines 37-55. Gillick’s phoneme-in-context represents the Markov Model used to represent inter-word context. It would have been obvious for a person having ordinary skill in the pertinent art, at the time the invention was made, that “transcription” as claimed is obvious to affect recognition because this is the way linguists have prepared speech models such as those used by Kao and Gillick for speech recognition. Using symbols shown by Gillick in the system of Kao would have been obvious because Kao teaches that the use of such a corpus is well known to implement transcription.

It is noted that some of the specific categories and process symbols taught in the applicant’s disclosure (i.e. figures 8(a)-(b)) are not explicitly taught in the prior art of record. However, the applicant claims some broadly as “acoustic-phonetic symbols and connectors” and some of the representations are specifically taught or rendered obvious by Kao’s classification examples in column 2, lines 30-40: “Nasalization” [applicant, figure 8(b)] = his nasal; “voicing” [Fig. 8(b)] = voiced/unvoiced; “Frication” [Fig. 8(b)] = fricative. Thus, the examples given by Kao show that it is obvious to utilize symbols not only for the phones but for other properties of the phones and how they interconnect. Further symbols are shown by Gillick in columns 28-29. It would have been inherent to use any of these symbols for transcription because the definition of “transcription” is the conversion of data from one language, code, medium to another, including reading, translating, and recording functions. Thus, the representations of these symbols as ink on paper is a transcription. The further details regarding decision trees and the training of models are for the purpose of further enhancing the otherwise well known phoneme

transcriptions to create more accurate representations thereof to perform speech recognition. Modeling stress, for example, is taught in col. 9, lines 48-50 using separate models.

[The change to claim 1 is addressed below with regard to claim 5. NOTE the current change incorporated limitations in canceled claim 6.]

Claim 2: Using the phone dataset that includes said speech data is taught by the training and mapping functions of figure 3.

Claim 3: Use in a speech recognizer is the purpose of the models.

Claim 4, 7: Phonetic dictionary is taught with the use of various phones used to recognize speech (see column 3). Use of the TIMIT database is taught by Kao in col. 2, lines 49-67.

Claims 5: Performing a “transformed phone dataset” is explicitly taught by Kao who clusters acoustically close triphones based on their context. The examples in column 2, indicate symbols representing similar sounding phones and the center phone of each triphone is therefore representative of the phone which an input sound corresponding thereto would be transcribed if properly recognized. A simple example would be recognizing a word with the letter (phone) d. While the dental placement of this phoneme might be labeled dh (i.e. – as sounded in the word “drive”). This could be compared to a glottal stop version of d (i.e. – as sounded in the word “hard”). The transcription of both versions of the d phoneme would still be the letter “d”. An explicit example of such a “transformation” be found in Gillick who teaches that it would be obvious to utilize a look up table from PIC (phoneme-in-context) to PEL (phonetic element) where speed is more important than memory storage in col. 24, lines 48-50.

Claims 9-19: The use of various symbols as transcription tools are obvious in view of the variety of text symbols used by Gillick. The applicant fails to teach any new, unobvious phonetic representation of speech. Choosing a particular character to represent old and well known representations of phone models fails to teach any unexpected result. Particular examples of the prior art include using and underline symbol to indicate left or right context as well as the symbolic indications of triphones with and without silence. The triphones themselves indicate composite-phones and there use with silence indicate beginning and end points (for example, left-silence indicates beginning of a word and right-silence indicates the end).

Claims 21-24: Transformation rules are taught by the use of triphones. Triphones represent a phone as it is influenced by left and right phones. When a match is made, the transcription of the phone is the center of the triphone. This process is performing a transformation of the input speech into the matched phones to form matched words. Such a transformation (or mapping) is explained by Gillick in columns 5-6 where he clearly teaches that his transformation (mapping) uses a decision tree. The decision tree was formed based on classification of the speech sounds, given its phonetic context in the given word or words. The rules that Kao and Gillick use for classification define the transformation rules (mapping) that is performed allowing one or more phones to be transformed (mapped) to a proper match based on the contextual rules. The only difference seems to be the applicant's reliance on symbols. However, it would have been obvious to utilize the symbols Kao shows in column 2 and that Gillick shows in columns 28 and 29 (see claims 8, 20 above), or combinations thereof, to represent the sounds they represent.

Claims 25-31, 33-43 and 45-50 are rejected under similar arguments as presented above.

[The changes to claim 25 is similar to that of claim 1 addressed above. The change to claim 25 added the limitations of canceled claim 30.]

The above rejection is repeated from the Final rejection mailed 26 March 2003 (paper #6) with added remarks placed in brackets []. The following remarks address newly added claims. However, the new claims that are rejected contain limitations which were clearly addressed in previous Final office action as noted (below).

Claim 51 has limitations which are the same as currently presented claim 1. The reasoning for rejecting these limitations was explained in paper #6 (previous Final Office action) as noted above.

Claim 53 has limitations which were addressed with respect to claims 14-18 regarding connector text symbols which may be placed to the left or right of phone symbols to indicate alternative sound contexts (stated in the previous Final Office action) as noted above.

Claim 55 has limitations which were addressed with respect to claims 21-24 (stated in the previous Final Office action) as noted above.

Claim 56 recitation of a “plurality of connectors that extend a corresponding one of said enhanced base phones” is a broader recitation of connectors as previously addressed with respect to claims 14-18. Applicant is directed to the reasoning for claims 14-18 noted above regarding connector text symbols which may be placed to the left or right of phone symbols to indicate alternative sound contexts.

The following remarks from the Final and Advisory actions (papers #6 and 8) are repeated below (items 5-11) and are part of this rejection. Reference to claims (6, 8 and 20) that are deleted or objected to have been removed.

REMARKS from Final action mailed 26 March 2003 (paper #8)

5. The applicant's arguments that the prior art fails to teach or suggest "enhanced phone set that includes acoustic-phonetic symbols and connectors for extending said enhanced phone set" is in error. Kao teaches in column 3, lines 5-7 that his corpus contains ...totally 34,629 utterances. The vocabulary size of the combined corpus is 16,165. The within word triphone count is 26,714. Because this combined corpus is larger than the TIMIT which has 5,222 utterances (col. 2, lines 53), this is considered enhanced. While Kao focuses on the ability to compute efficient models, it is clear that these models will allow the proper recognition and transcription of input speech into phones, words, sentences, etc. See, for example, column 8, lines 44-65 of Kao which indicates that most corpora transcriptions do not include inter-word pauses explicitly transcribed. Therefore, at the very least, his example grammar teaches an extended transcription symbol for handling a phone representing inter-word pauses and its connection with other phones and words.

More detailed explanations regarding claims 5 were added to the above rejection.

The disagreement about what the prior art teaches seems to be related to the applicant's use of symbols to represent speech. However, since the claims are directed towards "speech processing" the well known relationships between speech sounds and their representations as symbols must be taken into account. It is believed that the references utilized are sufficient to show that, in the context of speech recognition, it is well known to use combinations of symbols to represent sounds, sound classifications, and also combinations of sounds.

REMARKS (from Advisory Action – paper #8, mailed 7 July 2003)

6. According to the specification, “transforming” includes merge, split, replace or change in context as shown in figure 9. However, none of these particular “transformations” are claimed. Therefore, the claim language is deliberately broad so that it will encompass functional equivalents. The term transformation normally would read on something like a Fourier transform, for example. However, the applicant’s use indicates that the claim language would read on any model that defines how phones can go together or be separated rather than how any particular speech based parameters are mathematically calculated. Therefore, the prior art that shows different ways for dividing words into meaningful phones that can be put together in a variety of orders is an obvious example of such language.

The changes in context shown by Kao and explained with the example contexts of a <d> phone and the mapping of one context of phones into a single representation of a phonetic element both encompass the intended meaning under 35 US 112, sixth paragraph.

7. Both Kao and Gillick show how to develop phonetic models for speech recognition and it would have been obvious to combine the two because their teachings are very similar. For example, Kao teaches the use of triphones in col. 3, lines 42-48 and Gillick teaches a similar use of triphones in column 2, lines 47-53. Yet when these teachings are pointed out in rejecting canceled claim 6 (limitation now in claim 1), for example, the applicant inexplicably argues that such a combination of similar elements is not taught.

8. Many examples of the claim elements are provided. The applicant’s arguments are not understood because it is not believed that one of ordinary skill in the art would have such difficulty understanding the correspondence.

9. Both Kao and Gillick show transcription symbols (col. 2 and columns 28-29, respectively). Clearly, it would have been obvious to combine such similar elements because they are used for the same purpose in each reference to form models of speech. When this is

pointed out to the applicant, the applicant seems to ignore the teachings for what they represent to one of ordinary skill in the art.

10. The Examiner has given specific examples in rejecting the claims and can find no evidence of taking Official Notice.

11. The applicant's complaint about the Examiner's reference to a "TIMIT database" is not understood because the applicant has incorporated the SPHINX system by reference on page 8 of the specification. Kai-Fu Lee utilizes the TIMIT database in the SPHINX system. Therefore, the applicant teaches that his invention has a similar relationship as Kao because Kao makes reference to the TIMIT database in column 2 in teaching the known relationship between speech representations and phonetic transcriptions thereof. Thus, it would seem that the applicant is presenting contrary arguments by feigning a lack of understanding between phonetic dictionary and TIMIT while, at the same time, accusing the Examiner of failing to consider the claim terminology in context with the specification as required under 35 USC 112, sixth paragraph.

Allowable Subject Matter

12. Claims 8, 20, 32, 44, 52, 54 and 57-60 are objected to as depending upon a rejected base claim but would be allowable if re-written in independent form to include the limitations of any intervening dependent claims.

Claims 8, 32 and 52: The prior art does not teach "Articulator noise" defined on page 15 of the specification as 8 particular types of sounds.

Claims 20, 44 and 54: The prior art does not teach "epenthic vowel" defined on page 16 of the specification as a vowel added after a consonant for emphasis and is transcribed as symbol "a".

Claims 57-60: It is known to “merge adjacent phones”, “split”, “replace” and “change-in-context” but the prior art does not show the particular examples claimed.

13. Any response to this action should be mailed to:

Commissioner of Patents and Trademarks
Washington, D.C. 20231

or faxed to:

TC2600 Fax Center
(703) 872-9314

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA., Sixth Floor (Receptionist).

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to David D. Knepper whose telephone number is (703) 305-9644.

The examiner can normally be reached on Monday-Thursday from 07:30 a.m.-6:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil, can be reached on (703) 305-9645.

Any inquiry of a general nature or relating to the status of this application should be directed to customer service at (703) 306-0377.

The facsimile number for TC 2600 is (703) 872-9314.



David D. Knepper
Primary Examiner
Art Unit 2654